The accuracy of test day model evaluation for the Italian Holstein

The accuracy of test day model evaluation for the Italian Holstein

F. Canavesi[1], S. Biffani[1], G. Banos[2]

[1] Associazione Nazionale Allevatori Frisona Italiana. Cremona, Italy
[2] Aristotele University. Thessaloniki, Greece

Corresponding author: Fabiola Canavesi. Associazione Nazionale Allevatori Frisona Italiana, Ufficio Ricerca e Sviluppo. Via Bergamo 292, 26100 Cremona, Italy - Tel. +39 0372 474233 - Fax: +39 0392 474203 - Email: fabiolacanavesi@anafi.it

ABSTRACT: Genetic evaluation for production traits in the Holstein breed in Italy has been based on a Random Regression Test Day Model (RRTDM) since November 2004. More specifically, the model is a multiple lactation, multiple trait RRTDM, similar to the model used in Canada for official genetic evaluation.

Fixed regression curve effect include time, region, age at calving, parity and season of calving. Last changes in the model included a new definition of the proof scale and of the genetic base.

The accuracy of the model was assessed by analyzing residuals and testing Mendelian sampling trends. Residuals were normally distributed for all traits and had zero mean.

Residual trends for all the effects included in the model were analyzed also for effects not included in the model like number of milkings per day and number of days pregnant at the test date. Mendelian sampling did not show any significant trend over time both for cows and bulls.

Key words: Test day model, Accuracy, Residuals, Mendelian sampling.

INTRODUCTION – The work towards the development of the genetic evaluation for production traits and somatic cell score using a multiple trait multiple lactation Random Regression Test Day Model (RRTDM) based on Legendre polynomials with the same approach and programs of the Canadian Test Day Model (Jamrozik *et al.*, 1997; Jamrozik *et al.*, 1998; Schaeffer *et al.*, 1999; Kistemaker, 2003) started in 2001. The RRTDM has been official in Italy since November 2004.

From that day onward a big amount of time has been devoted to meeting farmers and industry people to explain the advantages of the new system and of all the additional information that can be used to better select the bulls that will help them increase their profit.

In the meantime research has started in order to improve the system and to address the many questions that users are raising while getting acquainted with the new system. One of the big differences between the old lactation Repeatability Animal Model and the RRTDM is that the latter assumes a more dynamic way of expressing genetic superiority, along the lactation and across lactations, that brings with itself more variation over time in bull proofs and this aspect is perhaps the most challenging for users.

As part of the work to improve the accuracy of the new model, the assessment of the level of accuracy of prediction of the current model was done at different levels.

Some checks are required by Interbull to verify the quality of national data that participate in the international bull genetic evaluation, others were performed in order to check the validity of the model, while new tools, like Mendelian sampling (Miglior *et al,* 2003) are being used and tested by different genetic evaluation units to evaluate the ability of the model to take into account selection effects.

The aim of this paper is to present some of the work that has been done on the Italian RRTDM in order to assess its accuracy; in particular, results will be presented for:

- Interbull Method III trend validation;
- residual analysis;
- Mendelian sampling trend.

MATERIAL AND METHODS – Data from the official evaluation of October 2006 were used for the analysis. The model is a multiple trait, multiple lactation random regression test day model for milk, fat , protein and somatic cell counts in the first three lactations as described in Muir et al. (2007).
The structure of the fixed effects curves that account for time, region, age, parity and season of calving effect (TRAPS) are as described in Canavesi *et al.* (2006). Data are preadjusted for heterogeneity of variance as described by Schaeffer *et al.* (2000). No pregnancy effect at the day of test is accounted for in the current structure.
Trend validation tests were run on October data following the methodology described by Boichard *et al.* (1995).
Residual effects were computed for each observation and their distribution was analyzed overall as well as separately by each fixed effect accounted for in the model. In addition, average residuals were computed for number of days of pregnancy at test date and number of milkings per day.
Mendelian sampling was investigated following a methodology suggested by Interbull For each animal, the Mendelian sampling deviation was computed as:

$$\hat{m}_a = \hat{u}_a - 0.5 \cdot (u_s + u_d)$$

where u_a, u_s and u_d are the predicted genetic merit for the animal, and its sire and dam, respectively. PEV of the mendelian sampling deviation was also computed and averaged across years using a software package called varMStol provided by Interbull.

RESULTS AND CONCLUSIONS – October 2006 official evaluation used 60,620,557 test day records, related to 4,398,555 cows calving between 1990 and 2006. The model evaluated a total of 5,276,591 animals , 456 TRAPS and a total of 4,733,651 herd test date parity levels.
Interbull method III validation did not show any significant effect due to including new information. The time effect was not significant for all traits; milk, fat , protein and somatic cell count.
Overall, residuals were normally distributed with an average around zero, as expected. Figures 1 and 2 report the average residual value for the effect of province and the effect of days of pregnancy which are not specifically included into the model. The province effect should be accounted for in the herd-test-day-parity effect and in the region effect and in the figure does not show any particular pattern. Provinces with less data do tend to have larger deviations from zero and this may be due to the low number of observations. The effect of days of pregnancy seems to be following a certain pattern, suggesting that it may have an impact on the traits of interest. Possible ways to include this effect in an improved model are currently being investigated.

Figure 1. Average residual values for milk kg by province code.

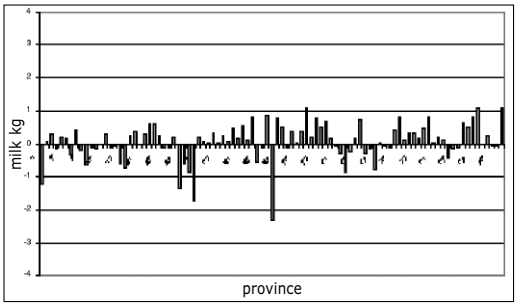
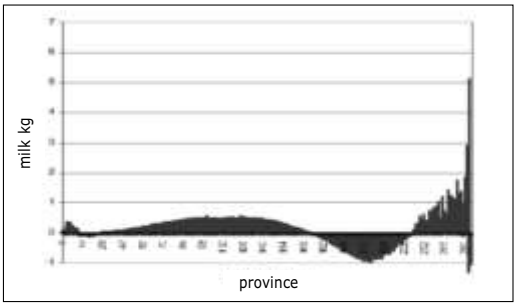


Figure 2. Average residual values for milk kg by number of days pregnancy at test date.



Mendelian sampling deviations did not seem to follow any precise pattern (Figures 3 and 4). The methodology proposed by Interbull did not show, at this stage of development, any serious problem in the current evaluation.
Although the current model does not seem to show any severe problem, work is ongoing in order to improve it.
A new model is being tested which includes the effect of each single year and the effect of number of days pregnancy at test date. The objective of the work is to further improve the overall accuracy of the system and its stability over time.

Figure 3. Mendelian sampling deviation for protein kg in first parity cows.

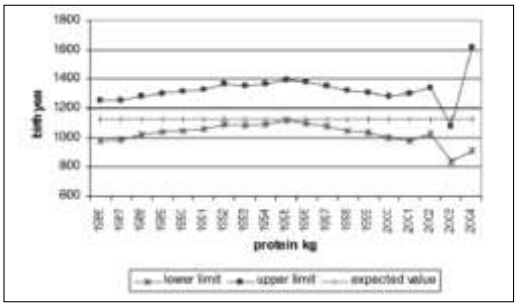
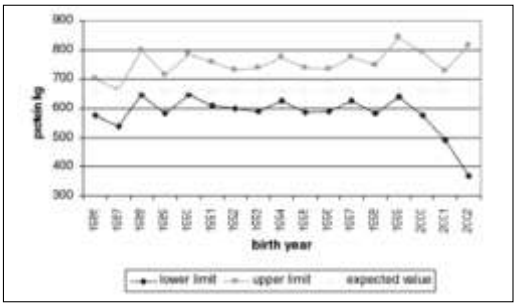


Figure 4. Mendelian sampling deviation for protein kg in Italian proven bulls.



REFERENCES – **Boichard,** D., B. Bonaiti, A. Barbat, S. Mattalia, 1995. Preliminary analysis of Mendelian sampling terms for genetic evaluation validation. J. Dairy Sci., 78: 431-437. **Canavesi,** F., S. Biffani, F. Biscarini, 2006. Italian test day evaluation: improving details that matter. Interbull bulletin 35: 17-21. **Jamrozik**, J., L.R. Schaeffer and F. Grignola. 1998. Genetic parameter for production traits and somatic cell score of Canadian Holsteins with a multiple trait random regression model. 6WCGALP 23:303-306. **Jamrozik**, J., L.R. Schaeffer, Z. Liu and G. Jansen. 1997. Multiple trait random regression test day model for production traits. Interbull Bulletin No. 16:43. **Kistemaker**, G.J., 2003. The Canadian Test Day Model using Legendre Polynomials. Interbull bulletin 31: 202-204. **Miglior,** F., P. Sullivan, B. Van Doormaal, 2003. Preliminary analysis of Mendelian sampling terms for genetic evaluation validation. Interbull Bulletin 29 : 183-187. **Muir**, B. L., G. Kistemaker , J. Jamrozik and F. Canavesi. 2007. Genetic Parameters for a Multiple-Trait Multiple-Lactation Random Regression Test Day Model in Italian Holsteins. J. Dairy Sci. (In press). **Schaeffer**, L.R., J. Jamrozik, G.J. Kistemaker, and B.J. Van Doormal. 2000. Experience with a test day model. J. Dairy Sci. 83: 1135-1144.